CLAIMS

1. A method comprising:

modeling an image with respect to multiple visual attentions to generate a respective set of attention objects (AOs) for each attention of the visual attentions;

analyzing the attention objects and corresponding attributes to optimize a rate of information gain as a function of information unit cost in terms of time associated with multiple image browsing modes; and

responsive to analyzing the attention objects, generating a browsing path to select ones of the attention objects, the browsing path being a trade off of time for space or space for time.

2. A method as recited in claim 1, wherein the multiple visual attentions are based on saliency, face, and text attention models.

3. A method as recited in claim 1, wherein the image browsing modes comprise a perusing mode and a skimming mode.

4. A method as recited in claim 1, wherein the select ones have relatively greater information fidelity as compared to different ones of the attention objects.

5. A method as recited in claim 1, wherein the corresponding attributes for each attention object of the AOs comprise a minimal perceptible time (MPT) for display of subject matter associated with the attention object.

6. A method as recited in claim 1, wherein the corresponding attributes for each attention object of the attention objects comprise a minimal perceptible time (MPT) for display of subject matter associated with the attention object, and wherein the method further comprises:

calculating the MPT as a function of:

- a number of words in the subject matter;
- whether the subject matter is for presentation to a viewer in a perusing image browsing mode or a skimming image browsing mode;
- user preferences; and/or,
- display context.

7. A method as recited in claim 1, wherein modeling further comprises:

creating a visual attention model for the image, the visual attention model being generated according to

$$\{AO_i\} = \{(ROI_i, AV_i, MPS_i, MPT_i)\}, \quad 1 \le i \le N \text{; and,}$$

wherein AO_i, represents an i^{th} AO within the image, ROI_i represents a region-of-interest of AO_i, AV_i represents an attention value of $AO_{i,}$, MPS_i represents a minimal perceptible size of AO_i, MPT_i represents a minimal perceptual time for display of subject matter associated with the AO_i, *and, N* represents a total number of AOs modeled from the image.

8. A method as recited in claim 1, wherein the browsing path comprises a number of successive path segments described as follows:

$$P = \{P_i\} = \{(SP_i, EP_i, SR_i, ER_i, T_i)\}, \quad 1 \le i \le N \text{, and}$$

wherein P_i represents an i^{th} path segment, SP_i represents a starting point of P_i, EP corresponds to an ending point of P_i , SR_i is a starting resolution of P_i , ER_i is an ending resolution of P_i, and T_i is a time cost for scrolling from SP_i to EP_i.

9. A method as recited in claim 1, wherein generating the browsing path further comprises creating the browsing path as a function of a fixation state or a shifting state.

10. A method as recited in claim 1, wherein generating the browsing path further comprises calculating an information fidelity for each AO, the information fidelity being a function of an attention value (AV) and a minimal perceptible time (MPT) for display of subject matter associated with the AO.

11. A method as recited in claim 1:

wherein each AO is a respective information block and wherein the image (I) is a modeled as a set of M x N evenly distributed information blocks I_{ij} as follows:

$$I = \{I_{ij}\} = \{(AV_{ij}, r_{ij})\}, \quad 1 \le i \le M, 1 \le j \le N, r_{ij} \in (0,1),$$

wherein (i, j) corresponds to a location at which the information block I_{ij} is modeled according to a visual attention, AV_{ij} is a visual attention value of I_{ij}, r_{ij} is the spatial scale of I_{ij}, representing the minimal spatial resolution to keep I_{ij} perceptible; and,

wherein generating the browsing path further comprises calculating an information fidelity (f_{RSVP}) for each AO, the information fidelity being calculated as follows for respective ones of the information blocks I_{ij}:

$$f_{RSVP}(I,T) = \int_0^T \sum_{I_{ij} \in I_{RSVP}(t)} AV_{ij} u(r(t) - r_{ij}) dt, I_{RSVP}(t) \subseteq I \text{; and}$$

wherein $I_{RSVP}(t)$ is a subset of the information blocks and varies with time and

$$r(t) = \max_{I_{ij} \in I_{RSVP}(t)} r_{ij} \le \min\left(\frac{Width_{Screen}}{Width_{I_{RSVP}(t)}}, \frac{Height_{Screen}}{Height_{I_{RSVP}(t)}}\right), \text{ which varies with space.}$$

12. A method as recited in claim 1, wherein generating the browsing path further comprises creating the browsing path in view of a skimming image-browsing mode as follows:

splitting one or more large AOs of the AOs into smaller AOs;

combining AOs in close proximity to one another into one or more attention groups;

arranging the attention groups in decreasing order based on respective attention values;

for each attention group of the attention groups:

selecting the attention group as a starting point;

calculating a total browsing time and information fidelity for each path of all possible paths from the starting point; and

if the total browsing time is greater than a browsing time threshold, discarding the path;

selecting a non-discarded path having a largest information fidelity as the browsing path, the browsing path connecting each of the attention groups.

13. A method as recited in claim 1, wherein generating the browsing path further comprises creating the browsing path in view of a perusing image-browsing mode as follows:

splitting one or more large AOs of the AOs into smaller AOs;

combining AOs in close proximity to one another into one or more attention groups;

arranging the attention groups in decreasing order based on respective attention values;

for each attention group of the attention groups:

selecting the attention group as a starting point;

for each path of all possible paths from the starting point:

calculating a total browsing time and an information fidelity; and

if the information fidelity is smaller than a browsing time threshold, discarding the path;

selecting a non-discarded path having a smallest browsing time as the browsing path, the browsing path connecting each of the attention groups.

14. A method as recited in claim 1, further comprising:

detecting user intervention during automatic playback of the browsing path;

responsive to detecting the user intervention:

recording all AOs S_r of the AOs that have not been browsed ;

identifying all AOs S_m of the AOs browsed during the user intervention;

regenerating the browsing path based on $S_r - S_m$; and

responsive to regenerating the browsing path and determining that there is at least a lull in user intervention, automatically navigating the browsing path.

15. A computer-readable medium comprising computer-program instructions executable by a processor for:

modeling an image with respect to multiple visual attentions to generate a respective set of attention objects (AOs) for each attention of the visual attentions;

analyzing the attention objects and corresponding attributes to optimize a rate of information gain as a function of information unit cost in terms of time associated with multiple image browsing modes; and

responsive to analyzing the attention objects, generating a browsing path to select ones of the attention objects, the browsing path being a trade off of time for space or space for time.

16. A computer-readable medium comprising computer-program instructions executable by a processor for:

modeling an image with respect to multiple visual attentions to generate a respective set of attention objects (AOs) for each attention of the visual attentions, the AOs representing respective regions of the image;

analyzing the attention objects and corresponding attributes to in view of a model for human browsing behavior, the model comprising fixation and shifting states, in the fixation state an interesting region of the regions is exploited for information, in the shifting state one region of the regions is replaced with another region of the regions as a function of image view manipulation operations such as scrolling or tabbing operations; and

responsive to the analyzing, optimizing a rate of information gain in terms of space as a function of information unit cost in terms of time associated with the model for human browsing behavior to generate a browsing path to select ones of the attention objects.

17. A computer-readable medium as recited in claim 16, wherein the multiple visual attentions are based on saliency, face, and text attention models.

18. A computer-readable medium as recited in claim 16, wherein the select ones have relatively greater information fidelity as compared to different ones of the attention objects.

19. A computer-readable medium as recited in claim 16, wherein the corresponding attributes for each attention object of the AOs comprise a minimal perceptible time (MPT) for display of subject matter associated with the attention object.

20. A computer-readable medium as recited in claim 16, wherein the computer-program instructions further comprise instructions for generating the browsing path as a number of successive path segments as follows:

$$P = \{P_i\} = \{(SP_i, EP_i, SR_i, ER_i, T_i)\}, \quad 1 \le i \le N, \text{ and}$$

wherein P_i represents an i^{th} path segment, SP_i represents a starting point of P_i, EP corresponds to an ending point of P_i, SR_i is a starting resolution of P_i, ER_i is an ending resolution of P_i, and T_i is a time cost for scrolling from SP_i to EP_i.

21. A computer-readable medium as recited in claim 16, wherein the computer-program for generating the browsing path further comprise instructions for calculating an information fidelity for each AO, the information fidelity being a function of an attention value (AV) and a minimal perceptible time (MPT) for display of subject matter associated with the AO.

22. A computer-readable medium as recited in claim 16, wherein each AO is a respective information block, and wherein the computer-program instructions further comprise instructions for:

representing the image (I) as a set of M x N evenly distributed information blocks I_{ij} as follows:

$$I = \{I_{ij}\} = \{(AV_{ij}, r_{ij})\}, \quad 1 \le i \le M, 1 \le j \le N, r_{ij} \in (0,1),$$

wherein (i, j) corresponds to a location at which the information block I_{ij} is modeled according to a visual attention, AV_{ij} is a visual attention value of I_{ij}, r_{ij} is the spatial scale of I_{ij}, representing the minimal spatial resolution to keep I_{ij} perceptible; and,

generating the browsing path further by calculating an information fidelity (f_{RSVP}) for each AO as follows for respective ones of the information blocks I_{ij}:

$$f_{RSVP}(I,T) = \int_0^T \sum_{I_{ij} \in I_{RSVP}(t)} AV_{ij} u(r(t) - r_{ij}) dt, I_{RSVP}(t) \subseteq I \text{; and}$$

wherein $I_{RSVP}(t)$ is a subset of the information blocks and varies with time and

$$r(t) = \max_{I_{ij} \in I_{RSVP}(t)} r_{ij} \le \min\left(\frac{Width_{Screen}}{Width_{I_{RSVP}(t)}}, \frac{Height_{Screen}}{Height_{I_{RSVP}(t)}}\right), \text{ which varies with space.}$$

23. A computer-readable medium as recited in claim 16, wherein the computer-program instructions for optimizing further comprise instructions for determining the rate of gain R as follows:

$$R = \frac{G}{T_B + T_w},$$

wherein, G is a total net amount of objectively valuable information gained as determined via information fidelity determinations, T_B is a total amount of time spent on shifting between subsequent fixation areas (AOs), T_W represents an exploiting cost, which is a total duration of the MPTs used while in a fixation state.

24. A computer-readable medium as recited in claim 23, wherein the computer-program instructions further comprise instructions for optimizing the rate of gain R either by maximizing information fidelity or by minimizing time cost.

25. A computer-readable medium as recited in claim 23, wherein the computer-program instructions further comprise instructions for optimizing the rate of gain R in the shifting mode as follows:

$$Given \quad T_P \le \lambda_T, \quad \underset{P}{Max}\left\{ f_P(I, T_P) \right\},$$

wherein, T_p represents a total amount of time spent for fixation and shifting in the browsing path P, λ_T is a threshold of maximal time cost, and I represents the image.

26. A computer-readable medium as recited in claim 23, wherein the computer-program instructions further comprise instructions for optimizing the rate of gain R in the shifting mode as follows:

$$Given \quad f_P(I,T_P) \geq \lambda_{AV}, \quad \underset{P}{Min}\{T_P\},$$

wherein, T_p represents a total amount of time spent for fixation and shifting in the browsing path P, λ_{AV} represents a minimal attention value or information percentage for attainment.

27. A computer-readable medium as recited in claim 16, further comprising:

detecting user intervention during automatic playback of the browsing path;

responsive to detecting the user intervention:

recording all AOs S_r of the AOs that have not been browsed ;

identifying all AOs S_m of the AOs browsed during the user intervention;

regenerating the browsing path based on $S_r - S_m$; and

responsive to regenerating the browsing path and determining that there is at least a lull in user intervention, automatically navigating the browsing path.

28. A computing device comprising a processor coupled to a memory, the memory comprising computer-program instructions executable by the processor for:

modeling an image with respect to multiple visual attentions to generate a respective set of attention objects (AOs) for each attention of the visual attentions, the AOs representing respective regions of the image;

analyzing the attention objects and corresponding attributes to in view of a model for human browsing behavior, the model comprising fixation and shifting states, in the fixation state an interesting region of the regions is exploited for information, in the shifting state one region of the regions is replaced with another region of the regions as a function of image view manipulation operations such as scrolling or tabbing operations; and

responsive to the analyzing, optimizing a rate of information gain in terms of space as a function of information unit cost in terms of time associated with the model for human browsing behavior to generate a browsing path to select ones of the attention objects.

29. A computing device as recited in claim 28, wherein the multiple visual attentions are based on saliency, face, and text attention models.

30. A computing device as recited in claim 28, wherein the select ones have relatively greater information fidelity as compared to different ones of the attention objects.

31. A computing device as recited in claim 28, wherein the corresponding attributes for each attention object of the AOs comprise a minimal perceptible time (MPT) for display of subject matter associated with the attention object.

32. A computing device as recited in claim 28, wherein the computer-program instructions further comprise instructions for generating the browsing path as a number of successive path segments as follows:

$$P = \{P_i\} = \{(SP_i, EP_i, SR_i, ER_i, T_i)\}, \quad 1 \le i \le N, \text{ and}$$

wherein P_i represents an i^{th} path segment, SP_i represents a starting point of P_i, EP corresponds to an ending point of P_i, SR_i is a starting resolution of P_i, ER_i is an ending resolution of P_i, and T_i is a time cost for scrolling from SP_i to EP_i.

33. A computing device as recited in claim 28, wherein the computer-program for generating the browsing path further comprise instructions for calculating an information fidelity for each AO, the information fidelity being a function of an attention value (AV) and a minimal perceptible time (MPT) for display of subject matter associated with the AO.

34. A computing device as recited in claim 28, wherein each AO is a respective information block, and wherein the computer-program instructions further comprise instructions for:

representing the image (*I*) as a set of *M* x *N* evenly distributed information blocks I_{ij} as follows:

$$I = \{I_{ij}\} = \{(AV_{ij}, r_{ij})\}, \quad 1 \le i \le M, 1 \le j \le N, r_{ij} \in (0,1),$$

wherein *(i, j)* corresponds to a location at which the information block I_{ij} is modeled according to a visual attention, AV_{ij} is a visual attention value of I_{ij}, r_{ij} is the spatial scale of I_{ij}, representing the minimal spatial resolution to keep I_{ij} perceptible; and,

generating the browsing path further by calculating an information fidelity (f_{RSVP}) for each AO as follows for respective ones of the information blocks I_{ij}:

$$f_{RSVP}(I,T) = \int_0^T \sum_{I_{ij} \in I_{RSVP}(t)} AV_{ij} u(r(t) - r_{ij}) dt, I_{RSVP}(t) \subseteq I \text{ ; and}$$

wherein $I_{RSVP}(t)$ is a subset of the information blocks and varies with time and

$$r(t) = \max_{I_{ij} \in I_{RSVP}(t)} r_{ij} \le \min\left(\frac{Width_{Screen}}{Width_{I_{RSVP}(t)}}, \frac{Height_{Screen}}{Height_{I_{RSVP}(t)}}\right), \text{ which varies with space.}$$

35. A computing device as recited in claim 28, wherein the computer-program instructions for optimizing further comprise instructions for determining the rate of gain R as follows:

$$R = \frac{G}{T_B + T_w},$$

wherein, G is a total net amount of objectively valuable information gained as determined via information fidelity determinations, T_B is a total amount of time spent on shifting between subsequent fixation areas (AOs), T_W represents an exploiting cost, which is a total duration of the MPTs used while in a fixation state.

36. A computing device as recited in claim 35, wherein the computer-program instructions further comprise instructions for optimizing the rate of gain R either by maximizing information fidelity or by minimizing time cost.

37. A computing device as recited in claim 35, wherein the computer-program instructions further comprise instructions for optimizing the rate of gain R in the shifting mode as follows:

$$\textit{Given} \quad T_P \leq \lambda_T, \quad \underset{P}{Max}\left\{f_p(I, T_P)\right\},$$

wherein, T_p represents a total amount of time spent for fixation and shifting in the browsing path P, λ_T is a threshold of maximal time cost, and I represents the image.

38. A computing device as recited in claim 35, wherein the computer-program instructions further comprise instructions for optimizing the rate of gain R in the shifting mode as follows:

$$Given \quad f_P(I,T_P) \geq \lambda_{AV}, \quad \underset{P}{Min}\{T_P\},$$

wherein, T_p represents a total amount of time spent for fixation and shifting in the browsing path P, λ_{AV} represents a minimal attention value or information percentage for attainment.

39. A computing device as recited in claim 28, further comprising:

detecting user intervention during automatic playback of the browsing path;

responsive to detecting the user intervention:

recording all AOs S_r of the AOs that have not been browsed ;

identifying all AOs S_m of the AOs browsed during the user intervention;

regenerating the browsing path based on $S_r - S_m$; and

responsive to regenerating the browsing path and determining that there is at least a lull in user intervention, automatically navigating the browsing path.

40. A computing device comprising:

means for determining respective attention value and image fidelity of attention objects derived from an image, specified minimal amounts of size and time conditions under which to present the attention objects to a user, and user-browsing characteristics;

means for adapting the image for presentation on a small display device to present objectively interesting portion(s) of the image based in part on display constraints of the device, and attributes of the attention objects; and

means for calculating a navigation path through the image that has been adapted for presentation on the small display device, the navigation path identifying a browsing path from one objectively important portion of the image to at least one other objectively important portion of the image, respective image portions being represented by individual ones of the attention objects, the browsing path being generated as a function of desired human browsing behavior associated with time and space.